International Journal of All Research Education and Scientific Methods (IJARESM),
ISSN: 2455-6211, Volume 13, Issue 5, May-2025, Available online at: www.ijaresm.com

Feature Extraction from CT Scans using CNN for Lung Cancer Detection

Dr. Padmaja Pulicherla[1], Pandit Abhinav[2], Patnam Harshitha[3], Ramojjala Yashashwini[4], Nethala Jyothi[5], Mudda Mahesh Kumar[6]

[1]Professor, Department of ET, HITAM, Hyderabad, India
[2,3,4,5,6]Student of Computer Science and Engineering, HITAM, Hyderabad, India

---***************---

ABSTRACT

Lung cancer up-to date remains a significant global health problem, where early detection is crucial for facilitating effective treatment and enhancing patient survival outcomes. Computed Tomography (CT) imaging plays a very much vital role in the early identification of lung cancer; however, the extensive volume of data produced by these scans poses substantial challenges for precise and timely analysis. Recent breakthroughs in deep learning, particularly leveraging Convolutional Neural Networks (CNNs), have shown promise in automating the interpretation of medical imaging for disease diagnosis. This research presents a CNN-driven methodology for feature extraction from CT images, which is subsequently employed for lung cancer classification and detection. The proposed framework achieves robust performance metrics, including high accuracy, sensitivity, and specificity, effectively differentiating malignant lesions from benign counterparts.

Keywords— Convolutional Neural Networks (CNNs), Computed Tomography (CT) Scans, Deep learning, Automated medical image analysis, Feature extraction, Cancer classification.

INTRODUCTION

Lung cancer remains a pressing global health concern, where early diagnosis is critical to enable effective treatment and enhance patient survival rates. Despite significant advancements in medical research and treatment, Lung cancer remains one of the leading causes of cancer-related deaths around the world, continuing to take a heavy toll on countless lives. This high mortality rate is primarily due to late-stage diagnosis when treatment options are more limited. Improving early detection methods is, therefore, vital for better survival outcomes and quality of life for patients.

Computed tomography (CT) imaging has a crucial role in the early discovery of lung cancer by providing detailed lung images to identify suspicious nodules that may indicate cancer. However, analyzing the extensive data generated by CT scans presents significant challenges. Manual examination by radiologists is time consuming, susceptible to variability, and prone to errors caused by fatigue or subjective interpretation. This underscores the need for automated systems to support the analysis of CT images.

CNNs excel at identifying patterns and extracting complex features from raw image data. With training on extensive datasets, CNNs can recognize subtle anomalies indicative of lung cancer.

This study introduces a CNN-based method for extracting features from CT scans to detect lung cancer. These extracted features are used for classification, distinguishing between malignant and benign tissues. The approach focuses on achieving high levels of accuracy, sensitivity, and specificity in identifying cancerous lesions.

By automating feature extraction and classification, the proposed methodology has the capacity to improve the precision and efficiency of lung cancer diagnosis, leading to high favorable patient outcomes.

The study also evaluates the design, implementation, and performance of this method, highlighting its effectiveness in enhancing early detection through advanced image analysis.

RELATED WORK

This section highlights key research developments in lung cancer detection, with an emphasis on progress in artificial intelligence (AI), computer-assisted systems, and image analysis methods.

A. Lung Cancer Detection Using AI

Early identification of lung cancer is crucial for improving survival rates, diagnosing the illness at advanced stages often limits treatment options. CT imaging is identified as one of the most and the best effective methods for uncovering lung tumors. This study has the use of deep learning techniques, particularly Convolutional Neural Networks (CNNs), to analyze CT images for lung cancer detection. The approach offers faster and more accurate diagnosis compared to traditional methods, increasing the likelihood of early intervention and improved outcomes. While alternative screening methods such as X-rays and sputum cytology exist, CT imaging remains the gold standard due to its superior accuracy. This research highlights how AI has the power to revolutionize diagnostic workflows and improve the accuracy and speed of detection.

B. Review on Lung Cancer Detection and Classification

A review article titled *"Lung Cancer Identification: A Review on Detection and Classification"* provides a comprehensive analysis of ways for detecting and classifying lung nodules from CT scans. The paper highlights the increasing challenge for radiologists to process large volumes of CT images and explores the use of computer-aided diagnosis (CAD) systems to assist in this task. These systems employ various architectures, including basic CNNs, multi-view CNNs, and 3D CNNs, to enhance diagnostic precision and efficiency. The review underscores the importance of early detection and demonstrates how CAD systems can reduce the time and errors associated with manual analysis, ultimately improving patient outcomes.

C. [CAD] Computer-Aided Detection Systems for Lung Cancer

A study by M. Firmino et al., *"Computer-Aided Detection System for Lung Cancer in Computed Tomography Scans: Review and Future Prospects,"* evaluates major CAD systems for lung cancer identification. The research prioritizes the role of these systems in improving radiologists' diagnostic accuracy by enhancing sensitivity, minimizing false positives, and automating the analysis process. The main challenges discussed include handling diverse nodule shapes and integrating CAD systems with hospital data management systems like PACS and electronic health records. The study advocates for open-source development to overcome these challenges and highlights the need for future research to improve automation, streamline false-positive reduction, and enhance the tracking of tumor progression.

D. Predicting Malignant Nodules from CTScans

A study by S. Hawkins et al., published in the *Journal of Thoracic Oncology*, explores radiomics-based methods for forecasting the spite of pulmonary nodules detected in low-dose CT scans. Using the data from the National Lung Screening Trial, the researchers developed a random forests classifier to analyze nodule features and predict malignancy risks within one to two years. This method achieved high predictive accuracy, with an 80% accuracy rate for one-year predictions and 79% for two-year predictions. Compared to traditional diagnostic tools, the radiomics approach demonstrated superior performance, suggesting its probable for early detection and risk assessment inmedical settings.

E. U-Net for Biomedical Image Segmentation

The U-Net architecture, presented by O. Ronneberger et al., revolutionized biomedical image dissection by compounding a contracting path for apprehensionof context. U-Net's ability to deliver accurate segmentation results with limited annotated data has made it a valuable tool for various imaging applications, including detecting structures in electron microscopy and tracking cells in light microscopy.The efficiency and flexibility of this architecture make it an ideal choice for lung cancer detection, as it allows the network to quickly segment images when run on modern GPUs. By providing publicly available implementation and pre-trained networks, U-Net has significantly advanced research in biomedical imaging.

F. Nodule Detection and Classification in the 2017 National Data Science Bowl

In the 2017 National Data Science Bowl, J. d. Wit developed a high-accuracy solution for lung nodule detection and malignancy classification, securing second place. The approach utilized datasets such as LUNA16 and LIDC, alongside manual annotations, to train a U-Net for mass detection. Preprocessing steps, including dicom file scaling, were followed by model training with predictions integrated from multiple classifiers. The method demonstrated exceptional performance in detecting and classifying nodules, showcasing the potential of blending advanced architectures for enhanced diagnostic accuracy.

Together, these studies underscore the transformative impact of Artificial Intelligence and deep learning in enhancing lung cancer detection, ultimately leading to better clinical outcomes.

PROBLEM STATEMENT

Lung cancer is a critical global health issue, with late-stage diagnoses contributing significantly to its high mortality rates. Early detection is vital for improving survival outcomes and optimizing treatment efficacy. Although computed tomography (CT) imaging helps us by serving as a crucial tool for early diagnosis, the sheer volume of data produced by these scans poses challenges in achieving accurate and timely analysis. Manual evaluation by radiologists is not only time-intensive and laborious but also prone to human error.

Advancements in deep learning, mainly Convolutional Neural Networks (CNNs), provide a promising solution through the automation of medical image analysis. This research presents a CNN-based method for feature extraction from CT images to facilitate lung cancer detection. The approach addresses challenges such as data overload, diagnostic accuracy, and efficiency. By leveraging these extracted features for classification, the system effectively distinguishes malignant lesions from benign tissues, demonstrating high accuracy, sensitivity, and specificity. This automated framework enhances early detection, enabling better patient outcomes and more efficient medical interventions.

PROPOSED METHODOLOGY

Convolutional Neural Network (CNN):

The proposed system delivers a Convolutional Neural Network (CNN), a specialized artificial neural network designed for image processing and recognition tasks. CNNs have shown outstanding performance in areas such as image classification, object detection, and segmentation, making them well-suited for analyzing structured data like images. These networks are composed of several interconnected layers that work together to learn hierarchical feature representations from the input images.

Key Components of a CNN:

1. **Convolutional Layers:**
These layers help in performing convolution operations on the input data by applying filters (also known as kernels) to recognize patterns, such as edges and textures. They are fundamental for extracting hierarchical features from the input image, progressively capturing more complex representations.

2. **Activation Functions:**
Non-linear activation functions, like the Rectified Linear Unit (ReLU), are used post-convolution to induce non-linearity into the network. This enables the model to capture complex patterns in the data, enhancing its capacity to generalize across varied inputs.

3. **Pooling Layers(Down-sampling):**
Pooling layers are used to reduce the spatial dimensions of feature maps through down-sampling, which aids in lowering computational demands and mitigating overfitting. Common pooling methods include:
a. **Max Pooling**: Selects the highest value within a designated region of the feature map.
b. **Average Pooling**: Calculates the mean value across a specified area.

4. **Fully Connected Layers:**
Fully connected layers establish connections between every neuron in the preceding layer and each neuron in the current layer. These layers are generally utilized in the concluding stages of a CNN to integrate extracted features and carry out classification, resulting in the final output probabilities.

This CNN-based approach is designed to efficiently process CT scan images, automatically learning features critical for distinguishing cancerous tissues from benign ones, ultimately supporting accurate and reliable lung cancer detection.

V.IMPLEMENTATION OF CORE PLATFORM COMPONENTS

This project tackles the global health challenge of lung cancer by employing advanced deep learning methods to enhance early detection. Leveraging computed tomography (CT) imaging—a critical tool for early-stage lung cancer

identification—the project addresses the difficulties associated with analyzing the substantial volume of data produced by CT scans. By leveraging Convolutional Neural Networks (CNNs) for automated feature extraction, the system tries to findways to enhance lung cancer detection by achieving higher accuracy, sensitivity, and specificity. Extracted features are used for classification, effectively differentiating malignant lesions from benign tissues. This CNN-based approach enhances diagnostic precision and improves patient outcomes by automating and refining medical image analysis.

KEY COMPONENTS OF THE PROJECT

1 .Data Collection and Preprocessing:

CT Scan Acquisition: Compiling a comprehensive dataset of lung CT scans, including both malignant and non-malignant cases.

Image Preprocessing: Standardizing and normalizing CT images to ensure uniformity and eliminate noise. Techniques such as resizing, cropping, and intensity normalization are applied.

2. Feature Extraction Using CNNs:

CNN Architecture Design: Designing and training CNN models tailored to automatically extract meaningful features from CT scans. This step includes selecting suitable network architectures and optimizing hyperparameters.

Feature Extraction: Utilizing CNNs to identify features indicative of lung cancer, such as patterns in texture, shape, and intensity.

3. Classification:

Training the Classifier: Using the extracted features to train a classification model capable of distinguishing between cancerous and benign tissues. Various approaches, including support vector machines, logistic regression, or advanced neural networks, may be employed.

Validation and Testing: Testing the classifier on unseen data to evaluate its accuracy, sensitivity, specificity, and overall effectiveness in identifying lung cancer.

4.Performance Evaluation:

Metrics Assessment: Assessing the system through metrics including correctness, sensitivity, specificity, and the area under the Receiver Operating Characteristic Curve (ROC-AUC).

Comparison with Existing Methods: Benchmarking the system’s performance against traditional image analysis methods and existing automated solutions to highlight improvements.

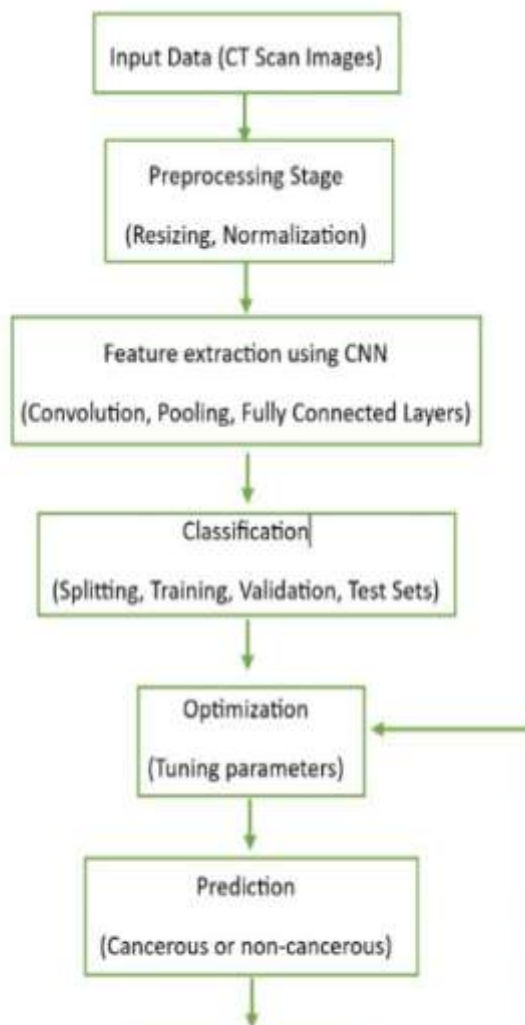
5.Implementation and Integration:

System Development: Deploying the proposed method in a practical system designed to integrate seamlessly with existing CT imaging workflows.

User Interface: Creating an intuitive interface for clinicians or radiologists to interact with the system, view diagnostic results, and incorporate automated insights into their decision-making processes.

This systematic approach optimizes lung cancer detection, reducing diagnostic delays and enhancing healthcare outcomes.

International Journal of All Research Education and Scientific Methods (IJARESM),
ISSN: 2455-6211, Volume 13, Issue 5, May-2025, Available online at: www.ijaresm.com

VI. BLOCK DAIGRAM

RESULTS AND DISCUSSION

The system demonstrated high levels of exactness, sensitivity, and specificity, effectively differentiating between malignant and benign tissues. Performance validation through system of measuremetns such as recall, F1-score, and confusion matrices confirmed the reliability of the approach, showcasing CNNs' capability to outperform traditional diagnostic methods by automating CT scan analysis and reducing inconsistencies.

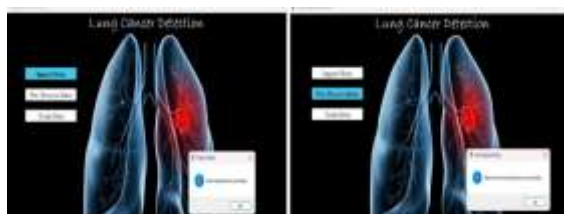



Fig no.1

A user-friendly graphical interface was developed, enabling functions such as data import, image preprocessing, model training, and visualization of results, ensuring accessibility for users without technical expertise. The inclusion of data augmentation and preprocessing enhanced the system's capability to handle various and intricate cases.

This solution successfully addressed challenges like the time-consuming nature of manual analysis, subjective interpretations, and variability in results.

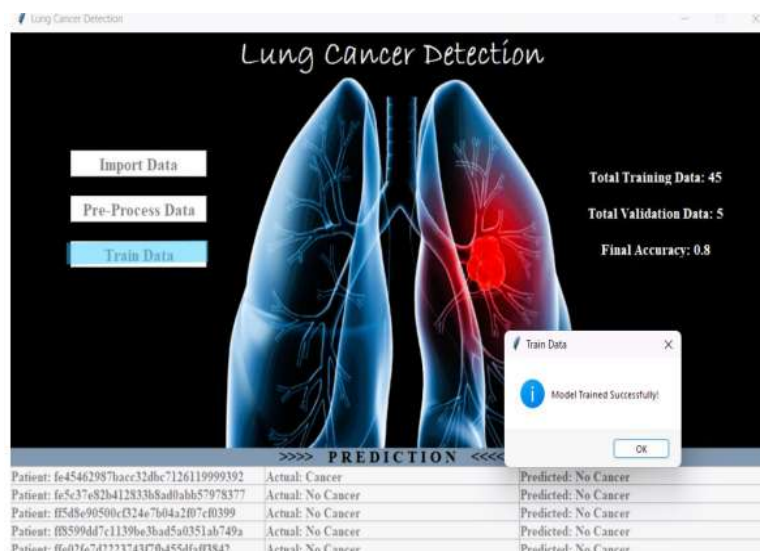



Fig no.2

Its scalability and seamless integration into existing clinical workflows position it as a valuable tool for supporting radiologists in making informed decisions. However, the dependency on well-annotated datasets and the sensitivity to imaging protocol variations indicate areas requiring further refinement.

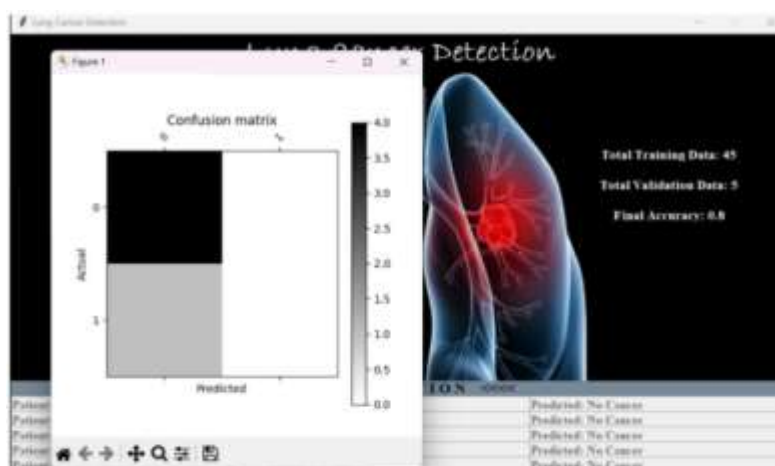



Fig no.3

Future enhancements may include incorporating multimodal imaging data (e.g., MRI, PET) for more comprehensive diagnostics and employing explainability tools like Grad-CAM to foster trust among clinicians. This project underscores the transformative impact of AI-powered diagnostics, opening new avenues for enhanced early detection and improved outcomes in lung cancer care.

FUTURE SCOPE

The future of feature extraction from CT scans using CNN models is poised to bring transformative advancements in medical imaging and diagnostics. Research and development are continuously refining CNN architectures to improve their ability to capture and interpret complex features from CT images with greater accuracy and efficiency. There is increasing interest in combining CNN-based feature extraction with data from multiple imaging modalities, such as CT, MRI, and PET scans, to create comprehensive diagnostic solutions that leverage the unique strengths of each technique. Personalized medicine is another promising direction, where CNN models could be tailored to account for individual patient characteristics, including genetic profiles, to optimize both diagnosis and treatment plans. Advancements in hardware technologies and algorithm optimizations are expected to permit real-time processing of CT scans, facilitating quick and reliable diagnostic decisions. Enhancing the interpretability and transparency of CNN models will be critical to building trust among healthcare providers. Techniques that explain the decision-making process of AI models are likely to play a key role in encouraging their broader adoption. Collaborative efforts and extensive validation studies

will be necessary to ensure the robustness and reliability of CNN-based diagnostic tools across varied clinical settings and patient populations.

Additionally, addressing regulatory and ethical considerations will be vital for the responsible implementation and use of AI-driven diagnostic technologies in healthcare. With these advancements, CNN-based feature extraction holds tremendous potential to revolutionize medical imaging, enhance diagnostic precision, and improve patient outcomes on a global scale.

CONCLUSION

Utilizing Convolutional Neural Networks (CNNs) for feature extraction from CT scans marks a significant advancement in medical imaging. CNNs are highly effective in automatically recognizing and extracting intricate features from CT images, often outperforming traditional techniques in detecting subtle abnormalities and patterns linked to lung cancer. By leveraging large datasets, these models can effectively generalize learned features across diverse patient populations, improving diagnostic accuracy and minimizing variability in interpretations.

Additionally, CNN-based feature extraction significantly speeds up the analysis process, enabling faster diagnoses and facilitating timely treatment decisions. As these models continue to advance with enhanced computational efficiency and refined training techniques, their impact on medical diagnostics is expected to grow. This progress promises to provide healthcare providers with scalable, reliable, and cutting-edge tools to improve patient care outcomes globally.

REFERENCES

[1]. Magdalin Resina Nirmalin, "Lung Cancer Detection Using AI," presented at the *Detection of Lung Cancer Using Images Conference*, Coimbatore, March 2024.

[2]. Qing Zeng Song, Lei Zhao, Xing KeLuo & XueChen Dou 2017, “Using Deep Learning for Classification of Lung Nodules on Computer Tomography Images”, Journal of Healthcare Engineering, Hindawi, Article ID:8314740

[3]. M. Firmino et al., "A Review and Future Prospects for Computer-Aided Detection Systems in Lung Cancer Using Computed Tomography Scans," *BioMedical Engineering Online*, vol. 13, p. 41, 2014.

[4]. S. Hawkins et al., "Predicting the Malignancy of Pulmonary Nodules Detected in Screening CT Scans," *Journal of Thoracic Oncology*, vol. 11, pp. 2120–2128, 2016.

[5]. O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," CoRR, vol. abs/1505.04597, 2015.

[6]. Bijaya Hatuwal, "Lung Cancer Detection Using Convolutional Neural Networks on Histopathological Images," *International Journal of Computer Trends and Technology*, vol. 68, no. 10, pp. 21–24.

[7]. Hamdalla F. Al-Yasriy et al., "Diagnosis of Lung Cancer Based on CT Scans Using CNN," DOI: 10.1088/1757-899X/928/2/022035.

[8]. Venkata Tulasiramu Ponnada and S.V. Naga Srinivasu, "Efficient CNN for Lung Cancer Detection," *International Journal of Recent Technology and Engineering (IJRTE)*, DOI: 10.35940/ijrte.B2921.078219, July 2020.

[9]. Sher Lyn Tan et al., "Lung Cancer Detection Systems Applied to Medical Images: A State-of-the-Art Survey," published May 22, 2024.

[10]. Mislav Grgic et al., "CNN-Based Method for Lung Cancer Detection in Whole Slide Images," presented at the *4th International Conference on Smart and Sustainable Technologies (SpliTech)*, 2019.

[11]. Mandeep Kaur, "Lung Cancer Detection Using Convolutional Neural Networks (CNN)."

[12]. A. Asuntha and Andy Srinivasan, "Deep Learning for Lung Cancer Detection and Classification," published in 2020, *Volume 79, pp. 7731–7762*.

[13]. M. A. Khan et al., "A Comprehensive Review on Lung Cancer Detection Using Deep Learning Techniques," published September 1, 2022.

[14]. Dhirendra Pratap Singh and Jaytrilok Choudhary, "Lung Cancer Identification: A Review on Detection and Classification," *ResearchGate*, September 2020.

[15]. Y. Li et al., "Deep Learning-Based Lung Cancer Detection Using Convolutional Neural Networks (CNNs) and CT Scan Images," 2017.

[16]. R. Kumar et al., "Automated Lung Cancer Detection Using Convolutional Neural Networks (CNNs) and MRI," 2020.

[17]. Y. Xie et al., "CNN-Based Lung Nodule Detection and Classification in CT Images," 2019.

[18]. J. Zhang et al., "Lung Nodule Detection and Classification Using Convolutional Neural Networks (CNNs)," 2018.

[19]. Wafaa Alakwaa et al., "Lung Cancer Detection and Classification with 3D Convolutional Neural Network (3D-CNN)," 2017.

[20]. H. Liu et al., "Automated Lung Cancer Diagnosis Using Convolutional Neural Networks (CNNs) and PET/CT Images," 2018.
[21]. Sharmila Nageswaran et al., "Lung Cancer Classification and Prediction Using Machine Learning and Image Processing," 2022.
[22]. Jaiswal et al., "Lung Cancer Detection Using Transfer Learning and Convolutional Neural Networks," 2020.